KS/McConnell

Received in ERN
SEP 1 8 1973

E7.3-10.981
CR-133760

Bi-Monthly Progress Report

Reporting Period: 1 July - 31 August 1973

A. Title of Investigation:

Evaluation and Comparison of ERTS Measurements of Major Crops and Soil Associations for Selected Sites in the Central United States.

ERTS-A Proposal Number: SR 050

B. Principal Investigator: M. F. Baumgardner
GSFC Identification Number: UN630

C. Problems:

Computer compatible tapes have continued to arrive semi-regularly. Digital data for the 20 March 1973 pass over the Lubbock Regional Test Site have not been received. Excellent ground observation data and underflight data for that ERTS pass are available for this important season.

D. Progress during this two-month reporting period is described below:

1. Mollisols in the Central United States

Soils in the order Mollisol are very important agricultural soils for crop and range production in the Great Plains States. Delineation of these soils in land-use mapping is very useful in soil management and conservation. This study was conducted to determine if the multispectral reflectance from Mollisols is consistant in the states of North Dakota, Iowa, Kansas, and Texas.

The multispectral reflectance measurements from ERTS were analyzed by computer-implemented pattern recognition techniques to locate and characterize the test sites. The differences observed in the Mollisols between the test sites is related to the geographic distribution of the sites. Analysis of variance results indicate that these differences due to geographic separation are significant. The observed differences in the reflectance may be attributed primarily to climatic differences of the widely ranging geographic locations of the Mollisols examined in this study. Also, differences in site illumination and atmospheric conditions between sites may have contributed to the observed differences.

Implications of the above results tend to support the observations of other researchers that ground control information should be obtained from each region under examination if acceptable results are to be expected. Another factor should be that all data should be obtained under as near the same atmospheric conditions as possible.

2. Ground Observations for the Lubbock Regional Test Site

During the period 4 July to 6 July detailed ground observations were made over part of the Lubbock Regional Test Site. Because of limited time available for ground observations, three counties (Hale, Lubbock, and Lynn) were chosen for detailed study. Observations of surface features were made in two ways by (1) completing ground observation forms for selected fields, and (2) taking low altitude photos of the selected fields.

Three North-South transects across Hale, Lubbock, and Lynn Counties were chosen for making ground observations. Detailed observations were made at road intersections. These intersections were at an interval of 2 to 4 miles to give a well distributed set of ground observations covering the three counties. At each intersection, observations were recorded for 4 fields.

The second phase of the ground observation mission was a low altitude flight to obtain color photography of the selected intersections and other areas of interest. The flight covered the intersections in sequence. A tape recorder was used to make a log of the flight and the photos.

3. Inventory of Lynn County Rangeland

A large portion of Lynn County, Texas (approx. 37,000 hectares) is rangeland. Using computer-aided data processing on ERTS MSS data a classification was made which attempted to show the degree of mesquite invasion in the rangeland. The ERTS data were from the 18 June 1973 pass (scene number 1330-16531).

By comparing the classification with available ground truth, color and color-IR photography taken 10 September 1972 and 20 March 1973 by NASA, and oblique low altitude air photos taken 5 July 1973, the accuracy of the classification was assessed. Three classes of pasture were identified: (1) clear pasture, mostly grasses; (2) mixture of grasses and mesquite; and (3) areas of thick mesquite. In addition, one class, called "other", included lakes and agricultural areas.

It was found that differences in rangeland composition could be mapped from spectral data using computer-aided processing techniques. The T Bar Ranch, which surrounds the Double Lakes west of Tahoka, Texas, was chosen as a test area for the classification of rangeland. It was found that the classification was reasonably consistant with the low altitude oblique air photos. The following table summarizes the classification results of the T Bar Ranch and some of the surrounding agricultural land.

	Areas	Percentage
Agriculture and lakes	34494	45.9
Rangeland	40674	54.1
Total	75168	100.0
Clear pasture	25831	63.5
Mixture of grasses and mesquite	7285	17.9
Thick mesquite	7558	18.6
Total	40674	100.0

Two unexpected results were obtained in this investigation: (1) prairie dog towns in relatively clear pasture were classified as agriculture, and (2) dry lake beds were classified as bare soil (agriculture). Apparently prairie dog burrows in clear pasture compose enough of the surface area to be spectrally similar to cultivated agricultural land. Several large alkaline lakes, which were dry when the data were collected, were also spectrally similar to cultivated agricultural land. This was surprising because the lakes are highly reflective alkaline lake deposits while the soil is a brown or reddish brown sandy soil. In both of the above cases of misclassification, confusion can be minimized by using spatial data obtained from maps or air photos for making the final analysis.

4. Crop Identification in Hale and Lamb Counties, Texas

Work has begun on the identification of crops in Hale and Lamb Counties. ERTS computer compatible tapes data collected 6 July 1973 (scene number 1348-16523) is being used in conjunction with ground observations made by the LARS staff and the cooperating ground observers in the Lubbock Regional Test Site.

E. Significant Results: See Attachment A.

F. Publications and Papers: None

G. Recommendations: None

H. No further additions or changes are contemplated for the ERTS standing orders.

CCT's representing 25 scenes have been returned to NDPF. These tapes were of limited use in this project because (1) the data did not fall sufficiently within a test site, or (2) the sites had excessive cloud cover. The Scene ID's for the returned CCT's are as follows:

1018-17592	1205-16585	1294-16530
1096-16530	1274-16423	1294-16533
1114-16534	1274-16430	1294-16535
1150-16532	1276-16531	1294-16542
1152-17021	1276-16534	1295-16585
1187-16571	1277-16590	1295-16591
1187-16574	1277-16595	1295-17000
1203-16472	1284-15543	1296-17050
1204-16542		

I. ERTS Image Descriptor Form: See Attachment B.

J. Data Request Form:

No ERTS Data Request Forms were submitted to GSFC/NDPF during this reporting period.

5

Rangeland Classification in Lynn County, Texas

Discipline Category:

ERTS Frame Number: 1330-16531 obtained 18 June 1973.

Run Number: 73041000

Classification Serial Number: 323232915, Tape (130), File (10)

ERTS Contract Number: NAS5-21785

Introduction:

A large portion of Lynn County, Texas (approx. 82,000 acres) is rangeland. Using computer-aided data processing on ERTS MSS data a classification was made which attempted to show the degree of mesquite invasion in the rangeland. The ERTS data was collected on 18 June 1973 (scene number 1330-16531).

Procedure:

An unsupervised classification was run for an area around Tahoka Lake, Texas. Classes from the unsupervised were correlated with the available ground observations. Unsupervised classes which corresponded to major rangeland composition groups were used as a basis for a classification of most of Lynn County.

Results:

Three classes of pasture were identified: (1) clear pasture, mostly grasses; (2) mixture of grasses and mesquite; and (3) areas of thick mesquite. In addition, one class called "other" included lakes and agricultural areas.

It was found that differences in rangeland composition could be mapped from spectral data using computer-aided processing techniques. The T Bar Ranch, which surrounds the Double Lakes west of Tahoka, Texas, was chosen as a test area for the classification of rangeland. It was found that the classification was reasonably consistant with the low altitude oblique air photos.

Attachment B

ERTS IMAGE DESCRIPTOR FORM

(See Instructions on Back)

DATE _____

PRINCIPAL INVESTIGATOR M. F. Baumgardner

GSFC UN630

ORGANIZATION LARS/Purdue University

NDPF USE ONLY
ID _____
N _____
ID _____

PRODUCT ID (INCLUDE BAND AND PRODUCT)	FREQUENTLY USED DESCRIPTORS*			DESCRIPTORS
	Ag	Soils	Urban	
1044-16595	x	x	x	
1057-16323	x	x		
1079-16583	x	x	x	
1007-16563	x	x		
1296-17005	x	x		Lakes Clouds
1330-16524	x	x	x	Rivers Escarpment
1330-16531	x	x	x	Lakes Escarpment
1348-16523	x	x	x	Escarpment Rivers
1348-16525	x	x	x	Lakes Escarpment

*FOR DESCRIPTORS WHICH WILL OCCUR FREQUENTLY, WRITE THE DESCRIPTOR TERMS IN THESE COLUMN HEADING SPACES NOW AND USE A CHECK (✓) MARK IN THE APPROPRIATE PRODUCT ID LINES. (FOR OTHER DESCRIPTORS, WRITE THE TERM UNDER THE DESCRIPTORS COLUMN).

MAIL TO ERTS USER SERVICES
CODE 563
BLDG 23 ROOM E413
NASA GSFC
GREENBELT, MD. 20771
301-982-5403

GSFC 37-2 (7/72)

Table 4.--Per-field classification matrix based on data from 3 overflights 1/

Group	No. of fields	Percent fields correct	No. of samples	Percent sample correct	Cotton	Corn	Soybeans	Grass	Winter wheat	Odd	Not classified
Cotton	38	63.2	927	85.0	24	0	2	0	1	0	11
Corn	7	14.3	58	20.7	0	1	0	1	1	0	4
Soybeans	58	25.9	852	44.2	9	3	15	3	7	1	20
Grass	31	9.7	240	29.6	3	1	1	3	2	0	21
Winter wheat	5	40.0	85	56.5	1	0	0	1	2	0	1
Odd	4	50.0	55	80.0	0	0	1	0	0	2	1
Totals	143	32.9	2217	60.4	37	5	19	8	13	3	58

1/ August 26, 1972, MSS bands 4, 5, 7.
September 14, 1972, MSS bands 5, 7.
October 2, 1972, MSS bands 4, 5, 6, 7.

Table 4, shows 58 fields out of 143 fields were not classified. A large number of fields were not classified because they were too small. This resulted because the required number of pixels for the classifier exceeded the number of pixels within the defined field. This problem was not present when the point classifiers were used. Therefore, the per-field classification system is less efficient for a State like Missouri which has many small fields.

Temporal Overlay

The next analysis investigated the value of a temporal overlay of several ERTS passes. This particular data set was a temporal overlay of two ERTS passes stacked on a third. Each pass could be compared with the three passes together. However, there were 3 bad bands in the total of 12. Two poor quality bands were in the September 14 imagery and one poor quality band was in the August 26 imagery. This makes it difficult to compare the three dates since the number of bands is confounded with dates. Nevertheless, the

C.M.'s for each date are presented in Tables 5, 6, and 7. These tables can be compared to the 9 band-overlay of Table 1, since they are all equal prior probability models.

Table 5.--Classification matrix for August 26, 1972 based on MSS bands 4, 5, and 7

Group	No. of sample points	Percent correct	Number of samples classified into					
			Cotton	Corn	Soybean	Grass	Winter wheat	Odd
Cotton	927	60.7	563	92	108	63	58	43
Corn	58	56.9	2	33	0	7	11	5
Soybeans	852	15.3	57	72	130	245	322	26
Grass	240	45.4	32	41	26	109	29	3
Winter wheat	85	51.8	5	6	10	15	44	5
Odd	55	69.1	6	4	3	3	1	38
Totals	2217		665	248	277	442	465	120

Overall performance 41.4

Table 6.--Classification matrix for September 14, 1972 based on MSS bands 5 and 7

Group	No. of sample points	Percent correct	Number of samples classified into					
			Cotton	Corn	Soybean	Grass	Winter wheat	Odd
Cotton	927	71.4	662	44	36	47	116	22
Corn	58	34.5	12	20	6	9	2	9
Soybeans	852	28.9	184	62	246	132	210	18
Grass	240	44.6	43	21	45	107	22	2
Winter wheat	85	68.2	6	12	0	9	58	0
Odd	55	47.3	3	16	0	2	8	26
Totals	2217		910	175	333	306	416	77

Overall performance 50.5

Table 7.--Classification matrix for October 2, 1973 based on MSS bands 4, 5, 6, and 7

Group	No. of sample points	Percent correct	Number of samples classified into					
			Cotton	Corn	Soybean	Grass	Winter wheat	Odd
Cotton	927	66.7	618	35	30	149	46	49
Corn	58	37.9	21	22	4	4	0	7
Soybeans	852	20.8	142	46	177	141	275	71
Grass	240	42.5	58	9	23	102	32	16
Winter wheat	85	56.5	15	1	7	10	48	4
Odd	55	56.4	5	7	1	8	2	31
Totals	2217		860	120	242	414	403	178

Overall performance 45.0

The temporal overlay classification of Table 1 shows an overall performance of 58.4 percent as compared to 41.4 percent, 50.5 percent and 45.0 percent, respectively, for Tables 5, 6, and 7. Based on these comparisons, the temporal overlay does improve the classification. However, ground truth data can become more difficult to interpret in the temporal overlay tapes because of changes in land use from one date to the next. Thus, the time of year becomes very important in areas where double cropping is common or preparation of land follows each crop. One must try to decide what crop to call the field when using temporal overlays.

It was originally intended to compare classification matrices and mean vectors for the three months to see if they changed significantly. However, since 3 of the 12 bands were of poor quality, the test would be nearly meaningless. Some multivariate tests will be done later to compare $\mu_1^{(2)} = \mu_2^{(2)}$.

Independent Test Data

The last exercise completed was to estimate the C.M. for Missouri on independent test data. Since the number of fields and points within are small and the area covered is large, we need much training data to represent the total area. It didn't seem possible to divide the set into halves and still have enough training data. It was decided to use a jackknife procedure. This procedure has the advantage of giving unbiased estimates that are simple to calculate. The data were divided into three equal subgroups, two groups were used to train with and the third group was used as a test group. This was repeated three times, each time with a different group used as test data. These three tables are presented separately, then the three are combined and presented to give an unbiased estimate of the classification matrix where independent test data are used. By using independent data, it is hoped that the biased caused by using the same data for both training and testing would be eliminated. But the variance of each item in the latter tables may be somewhat higher than those in the previous tables since a smaller data set was used.

One cotton field of 27 points was not included in any of the three groups. So the total in Table 11 is 27 pixels smaller than the total of earlier tables. Table 11, is the matrix sum of Tables 8, 9, and 10.

Table 8.--Classification matrix using August 26, 1972 M.S.S bands 4, 5, and 7 with subgroups 2 and 3 as training data and subgroup 1 as test data

Group	No. of sample points	Percent correct	Number of samples classified into					
			Cotton	Corn	Soybean	Grass	Winter wheat	Odd
Cotton.....	479	56.2	269	11	129	36	30	4
Soybeans.....	138	45.7	35	6	63	17	12	5
Grass.....	66	34.8	15	7	15	23	4	2
Winter wheat..	44	4.5	0	1	30	11	2	0
Odd.....	24	25.0	1	3	9	2	3	6
Totals.....	751		320	28	246	89	51	17

Overall performance 48.3

Table 9.--Classification matrix using August 26, 1972 M.S.S. bands 4, 5, and 7 with subgroups 1 and 3 as training data and subgroup 2 as test data

Group	No. of sample points	Percent correct	Number of samples classified into					
			Cotton	Corn	Soybean	Grass	Winter wheat	Odd
Cotton.....	290	57.6	167	36	11	19	20	37
Corn.....	29	13.8	1	4	0	8	2	14
Soybeans.....	308	13.0	48	53	40	20	131	16
Grass.....	42	28.6	1	11	4	12	14	0
Winter wheat..	26	30.8	0	2	8	2	8	6
Odd.....	31	100.0	0	0	0	0	0	31
Totals.....			217	106	63	61	175	104

Overall performance 36.1

Table 10.--Classification matrix using August 26, 1972 M.S.S. bands 4, 5, and 7 with subgroups 1 and 2 as training data and subgroup 3 as test data

Group	No. of sample points	Percent correct	Number of samples classified into					
			Cotton	Corn	Soybean	Grass	Winter wheat	Odd
Cotton	131	47.3	62	22	1	22	22	2
Corn	29	41.4	3	12	2	5	7	0
Soybeans	406	2.0	6	29	8	137	224	2
Grass	132	43.2	20	27	0	57	27	1
Winter wheat	15	0.0	5	2	0	8	0	0
Totals	713		96	92	11	229	280	5

Overall performance 19.5

Table 11.--Classification matrix combining tables 8, 9, and 10

Group	No. of sample points	Percent correct	Number of samples classified into					
			Cotton	Corn	Soybean	Grass	Winter wheat	Odd
Cotton	900	55.3	498	69	141	77	72	43
Corn	58	27.6	4	16	2	13	9	14
Soybeans	852	13.0	89	88	111	174	367	23
Grass	240	28.3	36	45	19	92	45	3
Winter wheat	85	11.8	5	5	38	21	10	6
Odd	55	67.3	1	3	9	2	3	37
Totals	2190		633	226	320	379	506	126

Overall performance 34.6

Intended Analysis of Idaho

Segment and field location has been completed in Idaho. ERTS tapes are being reformatted and when this operation is complete, analysis will begin on the LARS terminal at the Goddard Space Flight Center. The Analysis will be much the same as that of Missouri.